THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 60,, No. 7, September 1981
Printed in U.S.A.

Special-Information-Tone Frequency Detection

By A. FEUER

(Manuscript received January 20, 1981)

The ability to distinguish recorded announcements from other call types is an essential part of the mechanized service-evaluation process. To do this, all announcements will have a special-information-tone prefix. A frequency detector—which is based on the correlation functions of the received signal—would be used to decide which announcement was triggered by a specific call attempt. This paper evaluates the performance of the frequency detector in the presence of additive noise and frequency shift induced by the announcement machine. The theoretical results, based on a calibration frequency, are very encouraging. To verify that use of this frequency is feasible in practice, an algorithm is proposed and its performance evaluated to show that it compares favorably to the theoretical one.

I. INTRODUCTION

As part of the process of evaluating the end-to-end performance in the telephone network, a sample of call attempts is evaluated and the attempts are classified into several categories, such as completed, busy, recorded announcements, etc. To mechanize this classification process, a machine must have the ability to distinguish between completed calls and recorded announcements. Current planning for the mechanized systems envisions the use of special-information-tone (SIT) prefixes which are to be attached to recorded announcements and can then be automatically recognized by the mechanized classifier (as well as alert the customer to the fact that he is listening to a recorded announcement). By choosing four distinct SITs, each representing a certain category of recorded announcements, the classifier will have the ability to distinguish between these categories as well as to recognize a recorded announcement in general.

The SIT is defined as a sequence of three consecutive tones. To get four distinct SITs five frequencies were chosen: $\underline{f}_1 < \underline{f}_2 < \overline{f}_1 < \overline{f}_2 < f_3$, and each SIT consists of $\underline{f}_i$, $\overline{f}_j$, f_3. The third tone has a fixed frequency and

will be used for calibration as is described later. The actual values of these frequencies are 904.5 Hz, 985.4 Hz, 1356.8 Hz, 1440.2 Hz, and 1758.5 Hz. (The choice of these frequencies is mainly the result of constraints imposed by the CCITT definition of special information tones and tone generation and detection considerations.) To recognize which of the possible SITs was received, the machine has to detect in the first tone whether it was $\underline{f_1}$ or $\underline{f_2}$ and in the second tone whether it was $\overline{f_1}$ or $\overline{f_2}$. This means that the classification process of the announcement categories is reduced to the two frequency-detection processes mentioned above. However, the planned direct recording of the SIT followed by the recorded announcement on the various announcement machines in the telephone network may introduce significant degradation into the reproduced SIT. In addition to additive noise, which is common to all signals in the network, frequency flutter and frequency shift of considerable effect on the reproduced tones may occur. The flutter effect, having an oscillatory nature, can be minimized by averaging properly the received data. In this paper, we report on our investigation of the combined effects of additive noise and frequency shift on the detection process. The detection scheme considered here involves use of correlation functions, and we have concluded that a reliable detection of SIT frequencies is possible provided that a certain level of signal-to-noise ratio is ensured and the available data is properly used. It should be pointed out that in order to carry out the analysis we assume that the additive noise is white. We recognize, however, that in reality a colored additive noise may have a dominating effect on the error bounds.

The structure of the paper is as follows. Section II presents the basic ideas of frequency detection via the correlation functions. In Section III, a white additive noise is considered to be present and the performance of the detector in this case as a function of signal-to-noise ratio is evaluated. The effects of frequency shift are introduced in Section IV and the use of the calibration frequency to eliminate these effects is considered. The performance of the frequency detector as a function of the signal-to-noise ratio is evaluated and the results of using the calibration frequency are compared to the results when it is not used; a significant improvement is observed.

The performance evaluations in Sections III and IV are based on using the likelihood-ratio test. This, because of the complexity of the expressions involved, is not practically feasible; however, these evaluations provide bounds on the performance of other schemes. In Section V, an algorithm for a detection process is presented. This algorithm is simple enough to be practical and its performance compares favorably to the theoretical one.

II. CORRELATION DETECTOR

The use of correlation functions of a received signal for detection purposes was devised as part of the overall classification process by J. E. Walls of Bell Laboratories.[1] Some of the principles of the classification process that are relevant to our analysis are briefly described here.

Let $r(t)$ be a received signal observed for a period of length T. The correlation functions are defined by

$$C_n = \frac{1}{T}\int_{T/2}^{T/2} r(t)r(t + n\Delta)dt \qquad n = 0, 1, 2 \cdots, \tag{1}$$

where $\Delta < T$ (practically we want Δ to be small relative to T).

We note that C_0 is in fact the signal's average power in this time interval.

Since we are interested in frequency detection, let us assume for a moment that the received signal is a pure frequency, namely

$$r(t) = \sin 2\pi ft$$

and derive expressions for the correlation functions in this case. Then,

$$C_n = \frac{1}{T}\int_{T/2}^{T/2} \sin 2\pi ft \cdot \sin 2\pi f(t + n\Delta)dt,$$

or carrying out the simple integration results in

$$C_n = \frac{\cos 2\pi fn\Delta}{2}\left[1 - \frac{\sin 4\pi fT}{4\pi fT}\right]. \tag{2}$$

The power term C_0 and the next two terms C_1, C_2, are used for the detection process. Since

$$C_0 = \frac{1}{2}\left[1 - \frac{\sin 4\pi ft}{4\pi ft}\right]$$

$$C_1 = C_0 \cos 2\pi f\Delta$$

$$C_2 = C_0 \cos 2(2\pi f\Delta),$$

if we normalize C_1 and C_2 by the power C_0 we get the following relationship

$$\left[\frac{C_2}{C_0}\right] = 2\left[\frac{C_1}{C_0}\right]^2 - 1. \tag{3}$$

Expression (3) means that all pure frequencies correspond to a point on a parabola in the $(C_1/C_0, C_2/C_0)$ plane. Clearly this parabola exists only for $-1 \leq C_1/C_0 \leq 1$ and $-1 \leqq C_2/C_0 \leq 1$ (see Fig. 1) and C_1/C_0,

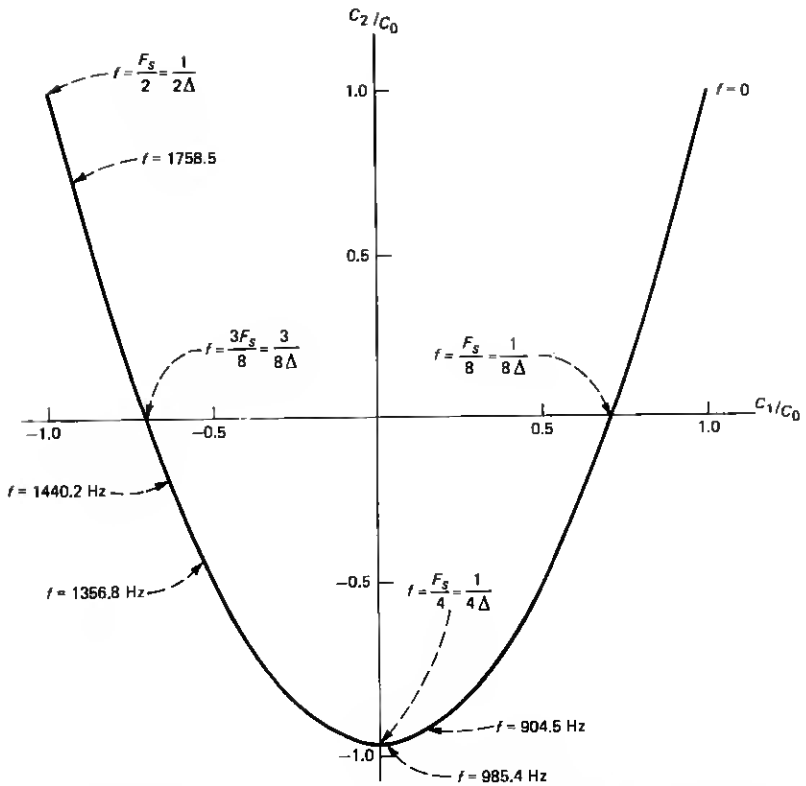


Fig. 1—Correspondence between various frequencies and the parabola on the $(C_1/C_0, C_2/C_0)$ plane.

C_2/C_0 are periodic functions of the frequency, so more than one frequency may correspond to the same point on the parabola. This means that for the detection process one must be concerned with the uniqueness of the points corresponding to the frequencies involved. For instance, if the two frequencies involved are $f_1 = \frac{1}{4}\,\Delta$ and $f_2 = \frac{3}{4}\,\Delta$, both correspond to the point $(0, -1)$ in the $(C_1/C_0, C_2/C_0)$ plane. This means that computing C_0, C_1, C_2 in this case is not sufficient to make the detection between these two frequencies possible even in the deterministic case (absence of noise).

However, the choice of the parameter $1/\Delta$ for a given set of frequencies involved in the detection process can ensure the necessary uniqueness. One way of doing it is to choose $1/\Delta$ to be larger than twice the largest frequency to be detected, and this will be sufficient, as can be seen in Fig. 1.

In the presence of noise, as is shown later, the point will shift from

the parabola towards the origin, while the origin itself corresponds to white noise for which both C_1 and C_2 are equal to zero.

In practice, instead of integration, the received signal is sampled and the sampled values are then used to get a good approximation of the correlation functions. Let $r(t)$ again denote the received signal, sampled with frequency $F_s = 1/\Delta$, and N the number of samples done in the observation interval T. Then eq. (1) will be replaced by

$$C_n = \frac{1}{N} \sum_{i=0}^{N-n-1} r_i r_{i+n}, \tag{4}$$

where

$$r_i = r\left[\frac{i}{F_s}\right].$$

Again, for a pure frequency, after some manipulations it can be shown that

$$C_n = \frac{1}{N} \sum_{i=0}^{N-n-1} \sin 2\pi \frac{f}{F_s} i \sin 2\pi \frac{f}{F_s} (i+n)$$

can be written as

$$C_n = \frac{1}{2} \cos n\, 2\pi \frac{f}{F_s} \left[\frac{N-n}{N} - \frac{1}{N} \cos(N-1)2\pi \frac{f}{F_s} \frac{\sin N\, 2\pi \frac{f}{F_s}}{\sin 2\pi \frac{f}{F_s}}\right]$$

$$+ \frac{1}{2} N \cos(N-1)2\pi \frac{f}{F_s} \frac{\cos N\, 2\pi \frac{f}{F_s} \sin n\, 2\pi \frac{f}{F_s}}{\sin 2\pi \frac{f}{F_s}}. \tag{5}$$

With the assumption that N is large compared to $n = 0, 1, 2$, we may write

$$C_n = \frac{1}{2} \cos 2\pi \frac{f}{F_s} n \left[1 - \frac{1}{N} \cos(N-1)2\pi \frac{f}{F_s} \frac{\sin N\, 2\pi \frac{f}{F_s}}{\sin 2\pi \frac{f}{F_s}}\right], \tag{6}$$

so that the relationship eq. (3) holds here as well. In the Appendix, the values of the correlation functions computed by eqs. (1) and (6) are given for the frequencies used for SIT. The error values introduced by going from eqs. (5) to (6) are also given in the Appendix for the same frequencies.

Suppose now that a frequency, one of possible two, is transmitted.

If on the receiving end we make sure that the sampling frequency is large enough compared to the two possible frequencies, we ensure the uniqueness of the correspondence between these frequencies and the parabola defined by eq. (3). With this assurance, we can sample the received signal, compute C_0, C_1, C_2 according to eq. (4), and then use the values C_1/C_0 and C_2/C_0 to decide which frequency was initially sent.

III. ADDITIVE NOISE

In the previous section, frequency detection using correlation functions was described. As long as the received signal contains nothing but the pure frequency originally sent, the problem is clearly deterministic. However, in practice there are various degradations that affect the received signal. In this section, we are interested in evaluating the performance of the correlation detector in the presence of additive white noise.

Thus, let the received signal be

$$r(t) = A \sin 2\pi f_k t + n(t) \qquad k = 1, 2,$$

where A corresponds to the signal-to-noise ratio and $n(t)$ is a normalized white Gaussian noise. The sampled values are

$$r_i = A \sin \theta_k \, i + n_i, \tag{7}$$

where

$$r_i = r\left[\frac{i}{F_s}\right], \; n_i = n\left[\frac{i}{F_s}\right], \; \theta_k = 2\pi \frac{f_k}{F_s},$$

$$k = 1, 2, \qquad i = 0, 1, 2, \cdots, N-1,$$

and n_i are independent identically distributed (IID) variables with zero mean and variance one.

Substitution of eq. (7) into eq. (4) results in

$$C_n^k = \frac{1}{N}\left[\sum_{i=0}^{N-n-1} A^2 \sin i\theta_k \sin(i+n)\theta_k + \sum_{i=0}^{N-n-1} A[n_{i+n}\sin i\theta + n_i \sin(i+n)\theta] + \sum_{i=0}^{N-n-1} n_i n_{i+n}\right]. \tag{8}$$

Since eq. (8) contains random variables, C_0, C_1, C_2, and hence, the ratios C_1/C_0 and C_2/C_0 are also random variables with certain density functions. Once these density functions are known the detection process becomes a standard problem described in detail in various textbooks on detection theory (see for example Ref. 2). We use the likelihood ratio test. To compute the threshold, we assume that the frequencies to be detected have equal a priori probabilities. The costs

are assumed to be zero and one for correct detection and error, respectively. With the knowledge of the density functions and the threshold, we can evaluate the performance of the detection process by computing the error probability.

The first step then, is to develop the expressions for the density functions of C_1/C_0 and C_2/C_0. However, in view of the complexity of eq. (8), rather than attempt to develop an exact expression, we will make use of a version of the Central Limit Theorem (Theorem 4.2.5 in Ref. 3) and develop approximate expressions for these density functions.

Let us now denote

$$S_n^k = \frac{A^2}{N} \sum_{i=0}^{N-n-1} \sin i\theta_k \sin(i+n)\theta_k. \tag{9a}$$

$$Y_{n1}^k = \frac{A}{N} \sum_{i=0}^{N-n-1} [n_{i+n} \sin i\theta_k + n_i \sin(i+n)\theta_k]. \tag{9b}$$

$$Y_{n2} = \frac{1}{N} \sum_{i=0}^{N-n-1} n_i n_{i+n}. \tag{9c}$$

Then,

$$C_n^k = S_n^k + Y_{n1}^k + Y_{n2}. \tag{10}$$

With some manipulation, we can get more convenient expressions for S_n^k and Y_{n1}^k:

$$S_n^k = \frac{A^2}{2} \cos n\theta_k \left[\frac{N-n}{N} - \frac{1}{N} \cos(N-1)\theta_k \frac{\sin N\theta_k}{\sin \theta_k} \right] + \frac{A^2}{2} N \cos(N-1)\theta_k \frac{\sin n\theta_k}{\sin \theta_k}.$$

$$Y_{n1}^k = 2\frac{A}{N} \cos n\theta_k \sum_{i=n}^{N-n-1} n_i \sin i\theta_k + \frac{A}{N} \left[\sum_{i=0}^{n-1} n_i \sin(i+n)\theta_k + \sum_{N-n}^{N-1} n_i \sin(i-n)\theta_k \right].$$

Now, since $N >> n$ we can write

$$S_n^k = \frac{A^2}{2} \cos n\theta_k \left[1 - \frac{1}{N} \cos(N-1)\theta_k \frac{\sin N\theta_k}{\sin \theta_k} \right]$$

and

$$Y_{n1}^k = 2\frac{A}{N} \cos n\theta_k \sum_{i=0}^{N-1} n_i \sin i\theta_k,$$

or as

$$S_0^k = \frac{A^2}{2}\left[1 - \frac{1}{N}\cos(N-1)\theta_k \frac{\sin N\theta_k}{\sin \theta_k}\right] \tag{11a}$$

$$Y_{01}^k = 2\frac{A}{N}\sum_{i=0}^{N-1} n_i \sin i\theta_k. \tag{11b}$$

We get

$$S_n^k = S_0^k \cos n\theta_k \tag{12a}$$

$$Y_{n1}^k = Y_{01}^k \cos n\theta_k. \tag{12b}$$

Denoting

$$x_1^k = \frac{C_1^k}{C_0^k}$$

$$x_2^k = \frac{C_2^k}{C_0^k}$$

and using eqs. (10) and (12) we get

$$x_1^k = \frac{[S_0^k + Y_{01}^k]\cos\theta_k + Y_{12}^k}{S_0^k + Y_{01}^k + Y_{02}^k} \tag{13a}$$

$$x_2^k = \frac{[S_0^k + Y_{01}^k]\cos 2\theta_k + Y_{22}^k}{S_0^k + Y_{01}^k + Y_{02}^k}. \tag{13b}$$

We observe now that x_1^k and x_2^k are functions of the random variables Y_{01}^k, Y_{02}^k, Y_{12}^k, Y_{22}^k, each one of which is a linear combination of IID random variables. The commonly used version of the Central Limit Theorem can be applied to conclude that the above four variables have Gaussian limiting distributions. This in turn enables us to use Theorem 4.2.5 in Ref. 3 to find the distributions of x_1^k and x_2^k.

The first step is to compute the means and variances of Y_{01}^k, Y_{02}^k, Y_{12}^k, and Y_{22}^k. Using the fact that the n_i's are IID Gaussian random variables with zero mean and variance one and using eqs. (9c) and (11b) we can readily verify that if

$$Y^k = \begin{bmatrix} Y_{01}^k \\ Y_{02}^k \\ Y_{12}^k \\ Y_{22}^k \end{bmatrix},$$

then

$$E\{Y^k\} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

and

$$\operatorname{cov}(Y^k) = E\left[[Y^k - E\{Y^k\}][Y^k - E\{Y^k\}]'\right] = \begin{bmatrix} \frac{4}{N} S_0^k & 0 & 0 & 0 \\ 0 & \frac{2}{N} & 0 & 0 \\ 0 & 0 & \frac{1}{N} & 0 \\ 0 & 0 & 0 & \frac{1}{N} \end{bmatrix}. \quad (14)$$

As in Ref. 3 denote vector b:

$$b = E[Y^k] = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

From eq. (14), the entries of Y^k, being Gaussian independent random variables, are jointly Gaussian, as is $\sqrt{N}(Y^k - b)$ with zero mean and covariance matrix

$$T = N \begin{bmatrix} \frac{4}{N} S_0^k & 0 & 0 & 0 \\ 0 & \frac{2}{N} & 0 & 0 \\ 0 & 0 & \frac{1}{N} & 0 \\ 0 & 0 & 0 & \frac{1}{N} \end{bmatrix}.$$

Now, let

$$x^k(Y^k) = \begin{bmatrix} x_1^k \\ x_2^k \end{bmatrix} \quad (15)$$

and

$$\phi_{ij} = \frac{1}{\sqrt{N}} \frac{\partial x_i^k}{\partial Y_j^k}\bigg|_b,$$

so by eq. (13)

$$\phi = \frac{1}{\sqrt{N}[S_0^k + 1]^2} \begin{bmatrix} \cos\theta_k & \cos 2\theta_k \\ -S_0^k \cos\theta_k & -S_0^k \cos 2\theta_k \\ S_0^k + 1 & 0 \\ 0 & S_0^k + 1 \end{bmatrix}.$$

All this establishes the preconditions for Theorem 4.2.5 in Ref. 3, which then states that the vector $[x^k(Y^k) - x^k(b)]$ has a Gaussian limiting distribution with zero mean and covariance matrix

$$R = \phi' T \phi. \tag{16}$$

Since N in our case is quite large, we may write

$$p_{x^k}(x^k) = p_{x|f_k}(x|f_k) = \frac{1}{2\pi |R_k|^{1/2}} \exp\left[-\frac{1}{2}[x - x^k(b)]' R_k^{-1} [x - x^k(b)] \right], \tag{17}$$

where

$$R_k = \frac{1}{N(S_0^k + 1)^4} \begin{bmatrix} (S_0^k)^2[2\cos^2\theta_k + 1] + 2S_0^k[2\cos^2\theta_k + 1] + 1, & 2S_0^k \cos\theta_k \cos 2\theta_k (S_0^k + 2), \\ 2S_0^k \cos\theta_k \cos 2\theta_k (S_0^k + 2), & (S_0^k)^2[2\cos^2 2\theta_k + 1] + 2S_0^k[2\cos^2 2\theta_k + 1] + 1 \end{bmatrix}$$

and

$$E\{x^k\} = x^k(b) = \frac{S_0^k}{S_0^k + 1} \begin{bmatrix} \cos\theta_k \\ \cos 2\theta_k \end{bmatrix}.$$

From eqs. (11a) and (17), we can readily observe the way the signal-to-noise ratio, A, affects the density function. If A gets very large, S_0^k becomes very large; then the mean value of x^k approaches the parabola and the entries of R_k become very small—namely, we approach the deterministic case described in the previous section. On the other hand, if A approaches zero, S_0^k goes to zero as well and the mean value approaches the origin. This is understandable since as A gets smaller, the effect of the noise gets larger, dominating the signal; and, being a white noise, the cross correlation functions C_1 and C_2 approach zero.

In Figs. 2 and 3, we see the effect of A as described above for two pairs of frequencies that were selected for the SIT. The sampling frequency is 4000 Hz. It should be noted that the mean values for each frequency are on a straight line as is obvious from eq. (17). In the figures, for every A, one equal-probability contour is drawn around the mean and the fact that these contours get smaller as A gets larger is expected since the entries of R_k are getting smaller as was pointed out earlier.

Once we have eq. (17) the detection is straightforward. With a priori probabilities and costs as described earlier, the measured data is used to compute the point in the $(C_1/C_0, C_2/C_0)$ plane and test which conditional density has higher value at this point. Then decide on the corresponding frequency as the one that was sent. More details of this procedure are described in Ref. 2. However, using the expressions we have for the density functions the process can be somewhat simplified. From eq. (17) and Ref. 2,

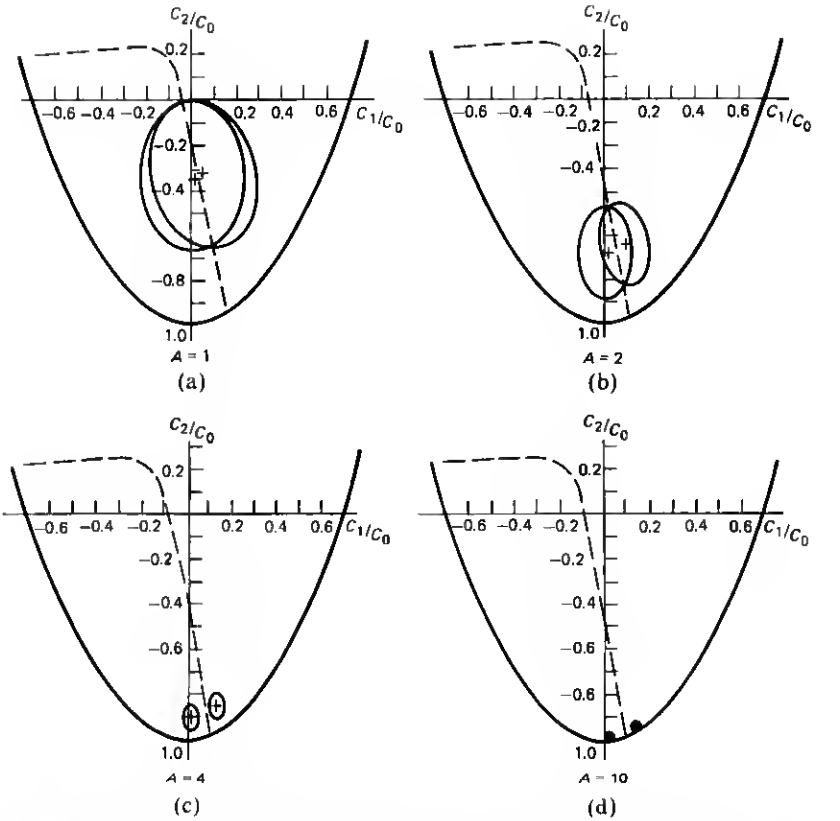


Fig. 2—Dependence of the density functions $P(x|f)$ on signal-to-noise ratio for 904.5 and 985.4 Hz (dashed lines are the equal likelihood points).

$$p_{x|f_1}(x|f_1) \underset{f_2}{\overset{f_1}{\gtrless}} p_{x|f_2}(x|f_2),$$

which means that if the inequality holds in one direction, we decide on f_1, and if in the other, f_2. This is equivalent to

$$1n|R_2| - 1n|R_1| + [x - x^2(b)]'R_2^{-1}[x - x^2(b)] - [x - x^1(b)]'R_1^{-1}[x - x^1(b)] \underset{f_2}{\overset{f_1}{\gtrless}} 0, \quad (18)$$

which with an equality sign is a line in the $(C_1/C_0, C_2/C_0)$ plane. The dotted lines in Figs. 2 and 3 correspond to tbe above-mentioned line. On one side of these lines one density function has higher values, on the other side the other density function. The detection then consists of deciding on which side of this line the measured point falls.

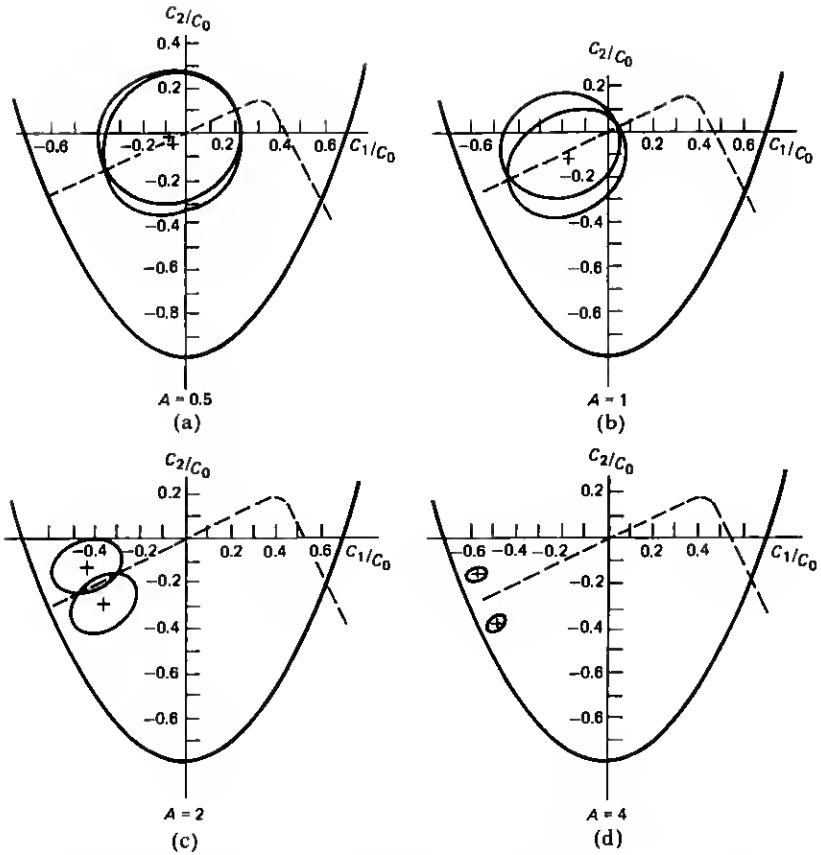


Fig. 3—Dependence of the density functions $p(x|f)$ on signal-to-noise ratio for 1356.8 and 1440.2 Hz (dashed lines are the equal likelihood points).

To evaluate the performance of this detector, we compute the error probability. Let B_k, $k = 1, 2$ be the part of the plane in which $p_{x|f_k}(x|f_k)$ has the larger value. Then the probability of error will be

$$\Pr(\epsilon) = \frac{1}{2}\left[\int_{B_2} p_{x|f_1}(x|f_1)dx + \int_{B_1} p_{x|f_2}(x|f_2)dx\right]. \tag{19}$$

Figure 4 shows the error probability as a function of A for, again, the two pairs of frequencies of interest in the SITS 904.5 Hz, 985.4 Hz, and 1356.8 Hz, 1440.2 Hz, and sampling frequency of 4000 Hz. Since, in general, signals transferred by tbe telephone network result in signal-to-noise ratios higber tban 8 dB, the results here are encouraging for both frequency pairs. It is interesting to note that the higher

frequency pair results in a better performance. Observing Figs. 2 and 3, one can see the reason for this difference in performance; the density functions for every signal-to-noise ratio are better separated in the higher pair. It turns out, that for every choice of a sampling frequency and difference in value, some pairs of frequencies—and not necessarily the lower valued frequencies—are more detectable than others.

IV. ADDITIVE NOISE AND FREQUENCY SHIFT

The use of announcement machines as tone generators will introduce two primary types of degradation, frequency flutter and frequency shift. In this section, we attempt to analyze the effects of the frequency shift with additive noise on the performance of the correlation detector. It is assumed that the flutter effect is eliminated by averaging a number of observations of each received tone.

The complexity of the expressions involved in the analysis here makes closed-form results very difficult if possible at all. For this reason, digital computer calculations were used as the main tool in the analysis.

Introducing the frequency shift, we get the expression for the received signal

$$r(t) = A \sin 2\pi \frac{(1 + n_d) f}{F_s} t + n(t),$$

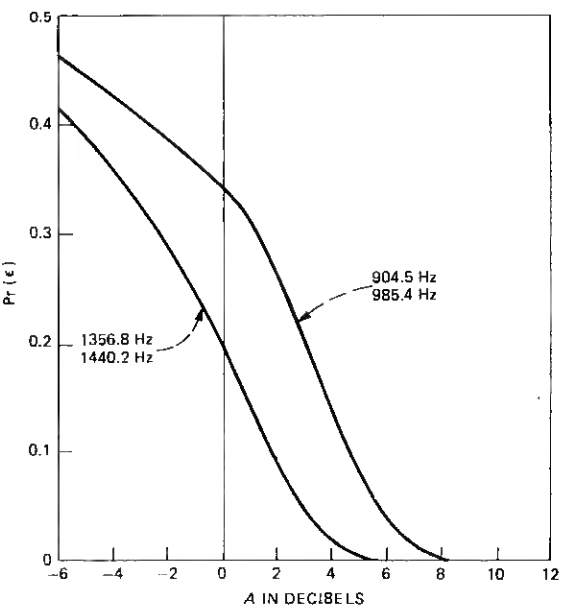


Fig. 4—Probability of error when no frequency shift is considered.

where n_d is assumed to be a random variable with extended beta distribution. This particular distribution is general enough to include many possibilities and agrees with the physical properties of n_d (namely $-1 \le n_d \le 1$). So n_d's assumed density function is

$$p_{n_d}(n_d) = \begin{cases} B(1+n_d)^{\alpha-1}(1-n_d)^{\beta-1} & \text{for } |n_d| \le 1 \\ 0 & \text{elsewhere} \end{cases},$$

where

$$B = \left|\frac{1}{2}\right|^{\alpha+\beta+1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

and $\alpha,\beta \ge 1$ are the two distribution parameters (in our computations we chose $\alpha = \beta = 10$ which fits reasonably the little data available for n_d).

The presence of this additional degradation causes the expressions developed in the previous section to be conditional on n_d. This means that now, rather than having an expression for $p_{x|f_k}(\cdot|\cdot)$, we have an expression for $p_{x|n_d,f_k}(\cdot|\cdot,\cdot)$ or using eq. (17) we may write

$$p_{x|n_d,f_k}(x|n_d,f_k) = \frac{1}{2\pi|R_k|^{1/2}} \exp\left[-\frac{1}{2}[x - E\{x^k\}]'R_k^{-1}[x - E\{x^k\}]\right], \quad (20)$$

where the expressions for R_k, $E\{x^k\}$, S_0^k are as before [see (11a) and (17)] and $\theta_k = 2\pi\dfrac{(1+n_d)f_k}{F_s}$.

Since for the likelihood ratio test $p_{x|f_k}(\cdot|\cdot)$ is needed, we can proceed to compute it using the relation

$$p_{x|f_k}(x|f_k) = \int_{-1}^{1} p_{x|n_d,f_k}(x|n_d,f_k)p_{n_d}(n_d)dn_d.$$

With this and eq. (19), the performance of the detector can be evaluated for this case where both additive noise and frequency shift are present. In Fig. 5 the probabilities of error in detection as a function of the signal-to-noise ratio are presented. Comparing these results to Fig. 4 makes it clear that the performance of the detector deteriorates very significantly when frequency shift is present. Even for a very high signal-to-noise ratio the frequency shift induces considerable error. The effect of the frequency shift alone, which provides lower bounds on the error probabilities in Fig. 5, can be calculated as follows. When A gets very large the received signal becomes a pure tone with a shifted frequency. This, however, implies that in this case it will be sufficient to consider only $x_1 = C_1/C_0$ for the detection. Since x_1 and n_d are

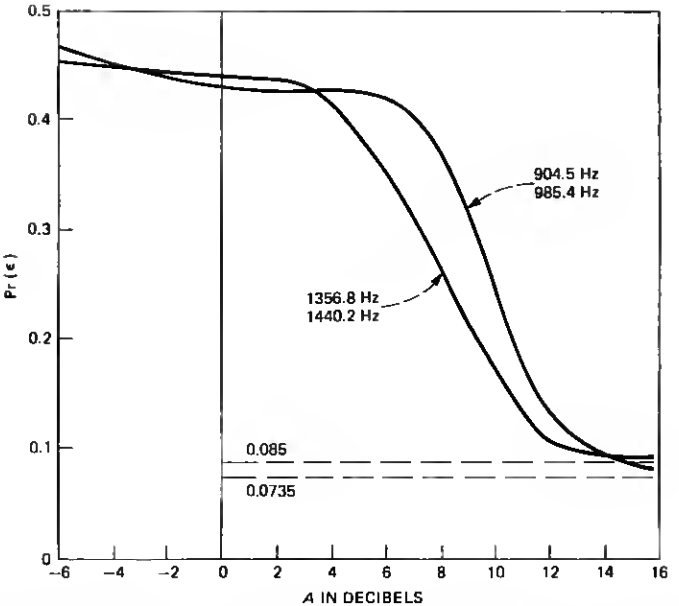


Fig. 5—Probability of error when frequency shift is considered but no calibration frequency is used.

related through

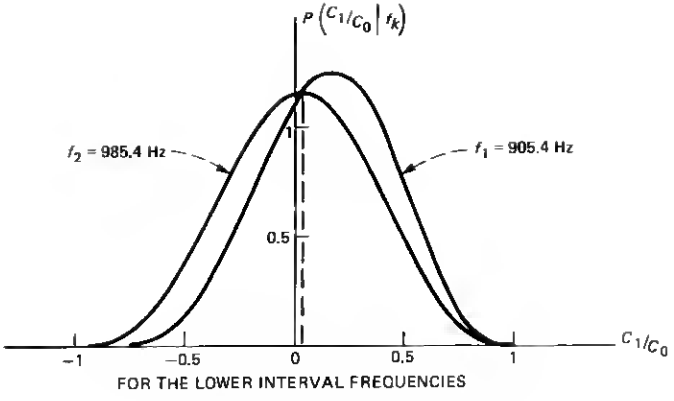
$$x_1 = \cos 2\pi(1 + n_d)\frac{f_k}{F_s},$$

the knowledge of $p_{n_d}(\cdot)$ makes the calculation of $p_{x_1|f_{ki}}(\cdot\,|\,\cdot)$ for each frequency f_k, straightforward. In Fig. 6 the density functions for the two pairs of frequencies are drawn and the thresholds for the detection in this case are the intersections of the density functions (also pointed out in the figure).

The technique to overcome the frequency shift is based on using the third tone for calibration in the detection of the first two tones. This means that a fixed tone is sent, processed in the receiving end, and the knowledge of its exact value and the corresponding measurements can be used to improve the detection of the first two tones.

To be more precise, let f_3 be the frequency of the third tone and $x^3 = [C_1^3/C_0^3,\ C_2^3/C_0^3]$ computed from the corresponding measured data. Then through eq. (20) we know that

$$p_{x^3|n_d f_3}(x^3\,|\,n_d, f_3) = \frac{1}{2\pi|R_3|^{1/2}} \exp\left[-\frac{1}{2}[x^3 - E\{x^3\}]'R_3^{-1}[x^3 - E\{x^3\}]\right],$$



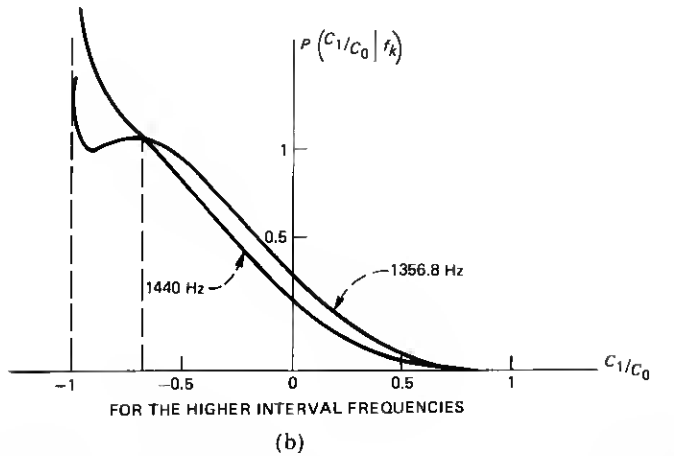


Fig. 6—Density functions required for detection when frequency shift is considered but the additive noise is ignored ($A \to \infty$).

where the expressions for $E\{x^3\}$ and R_3 are as in eqs. (11a) and (17) with f_3 substituted for f_k. Using this we can improve on our knowledge of the statistics of n_d by computing $p_{n_d|x^3 f_3}(.|.,.)$ through the relationship

$$p_{n_d|x^3 f_3}[n_d|x^3, f_3] = \frac{p_{x^3|n_d f_3}(x^3|n_d, f_3)p_{n_d}(n_d)}{\int_{-1}^{1} p_{x^3|n_d f_3}[x^3|n_d, f_3]p_{n_d}(n_d)dn_d}.$$

This improved data on n_d can then be used to calculate $p_{x|x^3 f_3 f_k}(.|.,.,.)$

$$p_{x|x^3 f_3 f_k}[x|x^3, f_k] = \int_{-1}^{1} p_{x|n_d f_k}(x|n_d, f_k)p_{n_d|x^3 f_3}[n_d|x^3, f_3]dn_d,$$

which in turn will be used for the detection process, namely

$$p_{x|x^3,f_3,f_1}[x|x^3, f_3, f_1] \underset{f_2}{\overset{f_1}{\gtrless}} p_{x|x^3,f_3,f_2}[x|x^3, f_3, f_2]. \tag{21}$$

In Figs. 7 and 8, we present—for this improved detection process with 1758.5 Hz as the calibration frequency—the error probabilities as a function of the signal-to-noise ratio. Note that the sampling frequency 4000 Hz is larger than 2×1758.5. Comparing this to the results without the calibration frequency, we observe considerable improvement, whereas the results computed with no frequency shift present provide a lower bound on error probabilities (or an upper bound on the improved detector's performance).

The case when $A \to \infty$ (i.e., when the additive noise becomes negligible) is again of special interest but very simple. Since now with the knowledge of both x_1^3 and f_3 the shift can be exactly computed,

$$x_1^3 = \cos 2\pi(1 + n_d)\frac{f_3}{F_s}$$

or

$$n_d = \frac{F_s}{2\pi f_3}\cos^{-1}x_1^3 - 1,$$

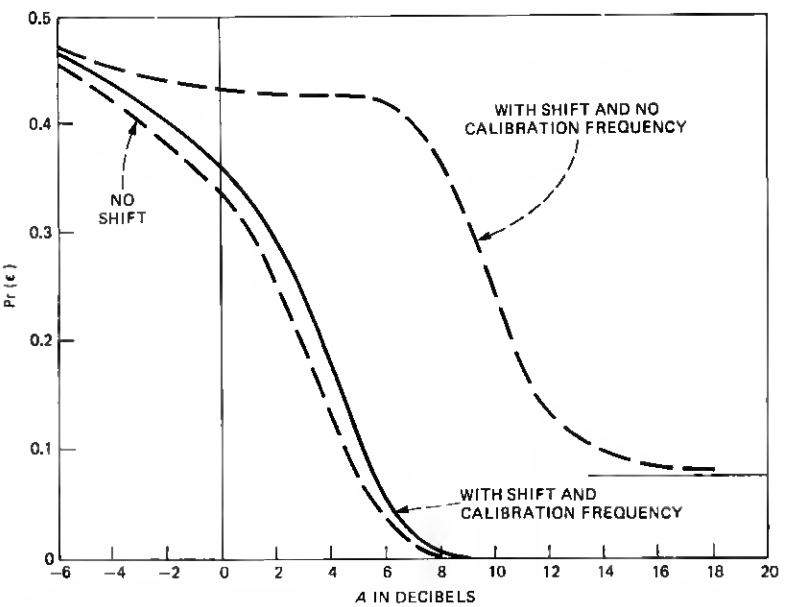


Fig. 7—Error probabilities with frequency shift and calibration frequency—lower interval frequencies.

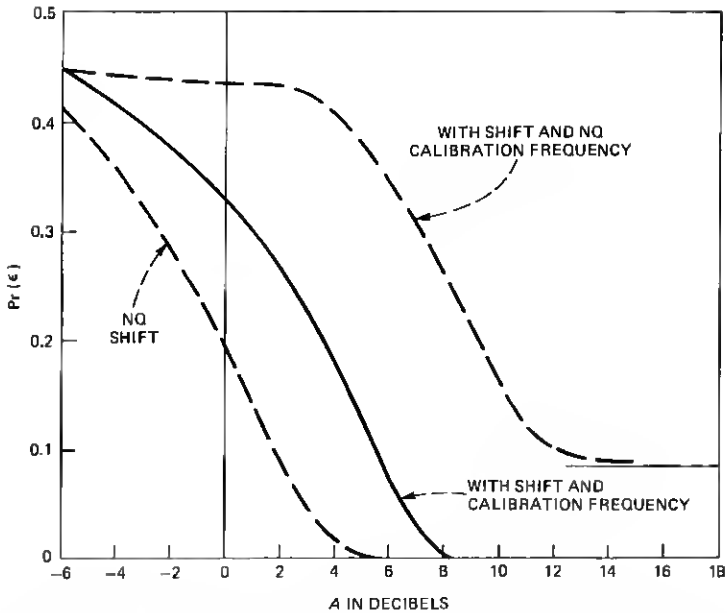


Fig. 8—Error probabilities with frequency shift and calibration frequency—higher interval frequencies.

and the detection of the first two tones becomes deterministic. The only source of problems in this case is the question of uniqueness of $\cos^{-1}x_1^3$. This can be overcome by considering both $\cos^{-1}x_1^3$ and $2\pi - \cos^{-1}x_1^3$ and with probability one, one of them will recover one of the two possible frequencies.

V. A SUGGESTED PRACTICAL CALIBRATION ALGORITHM

In the previous section, we described how the use of a calibration frequency can theoretically improve the detection of tones that are affected by frequency shift and additive noise. However, practical computation of the density functions required in eq. (21), or the separating line (where the two density functions are equal) is impossible. Hence, an algorithm is required that is both compatible to the theoretical exact density function separation and simple enough to be practically implemented. Such an algorithm is presented and its performance evaluated.

Let us first review the available data to be used for the detection. The received sequence of three tones is sampled and for each tone the correlation functions are calculated repeatedly and averaged to elimi-

nate frequency flutter effect. Then the cross-correlation functions are normalized to give the three vectors

$$x^1 = \begin{bmatrix} \frac{C_1^1}{C_0^1} \\ \frac{C_2^1}{C_0^1} \end{bmatrix} \quad x^2 = \begin{bmatrix} \frac{C_1^2}{C_0^2} \\ \frac{C_2^2}{C_0^2} \end{bmatrix} \quad x^3 = \begin{bmatrix} \frac{C_1^3}{C_0^3} \\ \frac{C_2^3}{C_0^3} \end{bmatrix}$$

corresponding to the three tones. Whereas the frequencies resulting in x^1 and x^2, respectively, are not known (in each case they can be either one of two possible frequencies), the third one is known to result from 1758.8 Hz. The idea is to use this knowledge to get an estimate of the frequency shift to help in the decision process of the first two tones.

The suggested algorithm (see Fig. 9) is as follows:

(*i*) Draw a line from the origin through x^3; it will intersect the parabola at

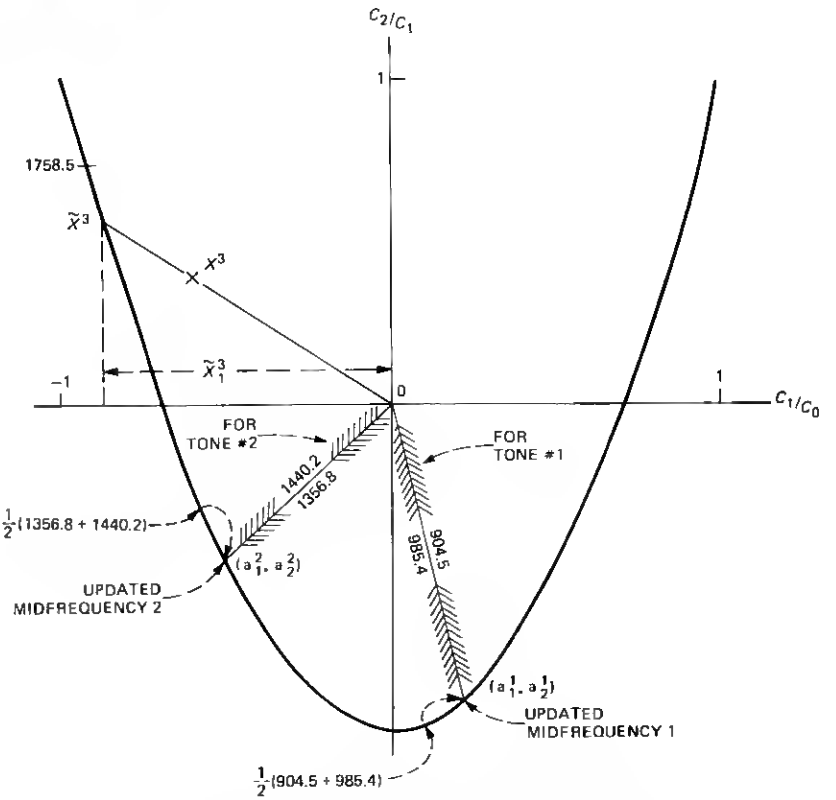


Fig. 9—Geometric interpretation of suggested algorithm.

$$\tilde{x}^3 = \begin{bmatrix} \tilde{x}_1^3 \\ \\ \tilde{x}_2^3 \end{bmatrix} = \begin{bmatrix} 1 \\ \\ \dfrac{x_2^3}{x_1^3} \end{bmatrix} \tilde{x}_1^3,$$

where

$$\tilde{x}_1^3 = \frac{x_2^3 + \sqrt{[x_2^3]^2 + 8[x_1^3]^2}}{2x_1^3}. \tag{22}$$

(*ii*) Use $\tilde{x}_1^3$ to compute the values

$$\left.\begin{aligned} a_1^i &= \cos\theta_i \\ a_2^i &= \cos 2\theta_i \end{aligned}\right\} i = 1, 2, \tag{23}$$

where

$$\theta_1 = \frac{904.5 + 985.4}{2 \times 1758.5} \cos^{-1}\tilde{x}_1^3$$

$$= 0.5374 \cos^{-1}\tilde{x}_1^3 \tag{24a}$$

and

$$\theta_2 = \frac{1356.8 + 1440.2}{2 \times 1758.5} \cos^{-1}\tilde{x}_1^3$$

$$= 0.7953 \cos^{-1}\tilde{x}_1^3. \tag{24b}$$

(*iii*) Draw the lines from the origin to the points (a_1^i, a_2^i): If x^i is counterclockwise away from the corresponding line, the ith tone is of the lower frequency, and if it is clockwise away, it is of the higher frequency.

The motivation behind this algorithm is quite simple. We have observed earlier that the additive noise effect is to shift the mean value of the pair $(C_1/C_0, C_2/C_0)$ towards the origin, whereas the frequency shift causes this pair to move along the parabola. The first step in the algorithm can be viewed then as isolation of the effect of the frequency shift. The point $\tilde{x}^3$ on the parabola is regarded as the result of the original frequency 1758.5 Hz, together with some shift that can now be estimated using the relationship

$$\tilde{x}_1^3 = \cos 2\pi \frac{1758.5}{4000} (1 + \hat{n}_d)$$

or

$$\hat{n}_d = \frac{4000}{2\pi\, 1758.5} \cos^{-1}\tilde{x}_1^3 - 1.$$

This estimate is then used to update the midfrequencies for each tone, $\frac{1}{2}(904.5 + 985.4)(1 + n_d)$ for the first and $\frac{1}{2}(1356.8 + 1440.2)(1 + n_d)$ for the second. The lines that connect the origin with the points

$$\begin{bmatrix} \cos\dfrac{\pi(904.5 + 985.4)}{4000}(1 + \hat{n}_d) \\ \cos\dfrac{2\pi(904.5 + 985.4)}{4000}(1 + \hat{n}_d) \end{bmatrix}$$

and

$$\begin{bmatrix} \cos\dfrac{\pi(1356.8 + 1440.2)}{4000}(1 + \hat{n}_d) \\ \cos\dfrac{2\pi(1356.8 + 1440.2)}{4000}(1 + \hat{n}_d) \end{bmatrix}$$

provide the threshold lines for the detection of the frequencies of the respective tones. In Fig. 9 the algorithm is described geometrically.

In Figs. 10 and 11 the performance of a detector using this algorithm is compared to the performance when exact separation is assumed. It is quite clear that the use of the algorithm results in a performance that is very close to the theoretical one.

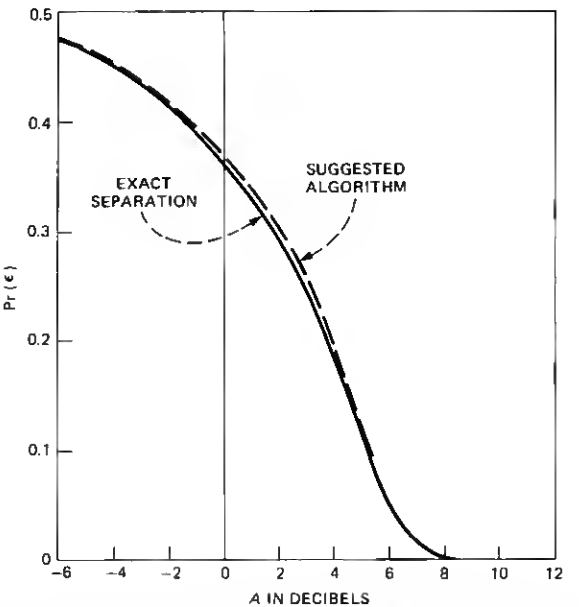


Fig. 10—Comparison of error probabilities for suggested detection algorithm to the exact separation—lower interval frequencies.

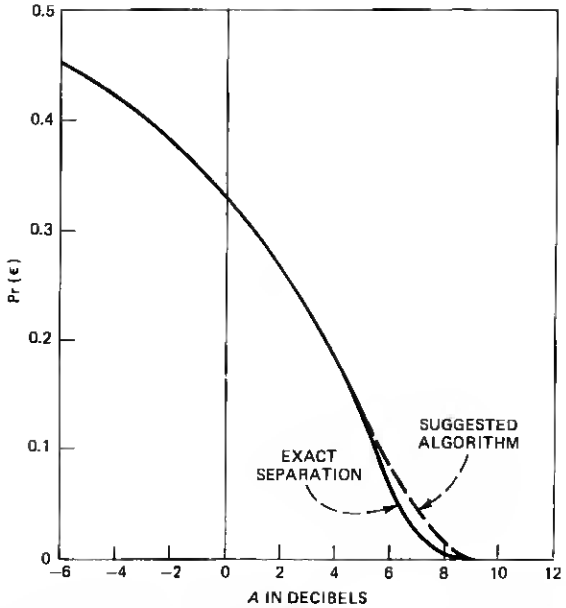


Fig. 11—Comparison of error probabilities for suggested detection algorithm to the exact separation—higher interval frequencies.

There are two difficulties in the described algorithm that will affect its performance. The first is because positive shifts above approximately 14 percent will result in frequencies higher than the critical one—$F_s/2$ (see Appendix)—for the third tone, and the points on the parabola are no longer uniquely related to their corresponding frequencies. This difficulty is inherent to the correlation function approach and can be eliminated only if a significantly higher sampling frequency is chosen (approximately 7000 Hz). However, if we assume that the shift is always less than the critical one, even for the proposed sampling frequency, the effect does not seem to be significant since in the model we have chosen for the frequency shift the probability of having shifts higher than 14 percent is quite low. The second difficulty, which is inherent in the proposed algorithm, arises when the resultant $\tilde{x}_1^3$ is less than -1. In this case, we propose simply to take it equal to -1 and again argue that the probability of this happening is very low even for small signal-to-noise ratios. Altogether, both difficulties, if treated as is suggested above, do not seem to affect the performance of the detection algorithm.

VI. CONCLUSION

In this paper, we addressed some of the problems in detecting recorded announcements encoded via the special information tones

(SIT). In particular, we discussed problems that arise when additive noise is present. We have assumed that the frequency flutter effects are eliminated by averaging several observations, and investigated in detail only the additive noise and frequency shift effects.

Our results support the conclusion that by properly using the information on the frequency shift, its effects can be made almost negligible and under these conditions high-performance SIT detection can be achieved.

The performance evaluations presented here make explicit use of a certain assumed model for the noise and frequency shift; however, the detection algorithm, which is proposed in Section V, is independent of any such assumptions. The performance of this algorithm is comparable to that theoretically achievable, and thus this algorithm is proposed for implementation in the SIT classification process.

VII. ACKNOWLEDGMENTS

The author gratefully acknowledges Mr. Chuck Antoniak for his useful comments and observations.

APPENDIX

Approximationa Uaed in Computing the Corralation Functions by Sampled Data.

f (Hz)	θ (rad)	CC_0	CC_1	CC_2	C_0	C_1	C_2	E_1	E_2
904.5	1.42	0.5	0.0747	−0.4778	0.4971	0.0743	−0.4749	0.0001	0.00004
985.4	1.55	0.499	0.0114	−0.4985	0.4982	0.0114	−0.4976	0.0008	0.00003
1356.8	2.13	0.4993	−0.2654	−0.2171	0.499	−0.2653	−0.217	0.0008	−0.0008
1440.2	2.26	0.4993	−0.3184	−0.0932	0.4997	−0.3186	−0.0933	0.0003	−0.0004
1758.5	2.76	0.5004	−0.4649	−0.3632	0.496	−0.4608	0.36	−0.0019	0.0034

Continuously computed correlation functions [see eq. (2)]:

$$CC_n = \frac{1}{2}\cos(n\theta)\left[1 - \frac{\sin 4\pi ft}{4\pi ft}\right]$$

Approximated correlation functions [see eq. (6)]:

$$C_n = \frac{1}{2}\cos(n\theta)\left\{1 - \frac{1}{N}\cos[(N-1)\theta]\frac{\sin(N\theta)}{\sin\theta}\right\}$$

Differences between discretely computed correlation functions [see eq. (5)] and their approximated values.

$$E_n = \frac{1}{2N}\cos[(N-1)\theta]\frac{\sin(n\theta)}{\sin\theta}\cos(N\theta).$$

[Note: $F_s = 4000$ Hz; $N = 167$; $T = 1/24$ seconds and $\theta = 2\pi\,(f/F_s)$].

REFERENCES

1. J. E. Walls, private communication.
2. L. H. Van Trees, *Detection, Estimation, and Modulation Theory*, Part 1, New York: John Wiley, 1968.
3. T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Canada: John Wiley, 1958.